ARMY RESEARCH LABORATORY

Nonparametric Methods for Multivariate Analysis Using Statistically Equivalent Blocks

Keith E. Matthews
U.S. MILITARY ACADEMY

Malcolm S. Taylor
U.S. ARMY RESEARCH LABORATORY

ARL-TR-992 April 1996

NOTICES

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1996	Final, June 1994 - September 1995

4. TITLE AND SUBTITLE
Nonparametric Methods for Multivariate Analysis Using Statistically Equivalent Blocks

5. FUNDING NUMBERS
JONO 6US400

6. AUTHOR(S)
Keith E. Matthews* and Malcolm S. Taylor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: AMSRL-SC-S
Aberdeen Proving Ground, MD 21005-5067

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-992

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

10.SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
*U.S. Military Academy, Department of Mathematical Sciences, West Point, NY 10996-1786

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Multivariate statistical procedures developed under normality assumptions are well advanced. Some of these procedures claim robustness properties, especially in a large sample situation, that may serve to broaden their range of application. Nonparametric methods for multivariate analysis have been pursued, but their more complete development awaits further research. This report considers nonparametric multivariate hypothesis testing in both one- and two-sample situations. Comparable univariate procedures do not extend readily to higher dimensions. The methods considered are based on the properties of statistically equivalent blocks. A new approach using a proximity-based cutting function for block construction is described. Statistically equivalent blocks, while holding the promise of important practical application, has received limited research attention.

14. SUBJECT TERMS
nonparametric statistics, multivariate analysis, goodness-of-fit, statistically equivalent blocks

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

INTENTIONALLY LEFT BLANK.

TABLE OF CONTENTS

INTENTIONALLY LEFT BLANK.

LIST OF FIGURES

LIST OF TABLES

INTENTIONALLY LEFT BLANK.

1. INTRODUCTION

Multivariate statistical procedures developed under normality assumptions are well advanced (see, for example, Anderson [1958] and Morrison [1976]). Some of these procedures claim robustness properties, especially in a large sample situation, that may serve to broaden their range of application. Nonparametric methods for multivariate analysis have been pursued, notably by Puri and Sen (1971), but their more complete development awaits further research.

This report considers multivariate hypothesis testing in both one-sample and two-sample situations. Comparable univariate procedures do not extend readily to higher dimensions. The methods considered are based on the properties of statistically equivalent blocks, which have received attention from a number of researchers, including Fraser (1957) in a tolerance interval context and Anderson (1966) and Wilks (1962) in an inferential setting.

In section 2 the mechanics of the procedure, along with the supporting mathematics, are given. In section 3 statistically equivalent blocks are applied in one- and two-sample situations. In section 4 proximity-based cutting functions are introduced and applied in the two-sample setting.

2. STATISTICALLY EQUIVALENT BLOCKS

The intent of the construction detailed in this section is to reduce the dimension of the problem in order to exploit traditional univariate methods. This is begun by partitioning the p-dimensional real product space $\mathbf{R}^p$, containing the observations into subspaces or blocks. The partition is effected through the use of functions $h:\mathbf{R}^p \to \mathbf{R}$ called cutting functions.

2.1 <u>Construction of Blocks</u>. Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be n observations of a p-component random vector $\mathbf{x}$ with distribution function $F(\mathbf{x})$ and let $h_1(\mathbf{x}), ..., h_n(\mathbf{x})$ be n (not necessarily distinct) real functions. The functions $h_i(\mathbf{x})$, i = 1, ..., n will be used to impose an order on the vectors $\mathbf{x}_1, ..., \mathbf{x}_n$. The value of the subscript i of the function $h_i(\mathbf{x})$ does not imply an order of application; i.e., $h_1(\mathbf{x})$ is not necessarily applied first, $h_2(\mathbf{x})$ second, etc. To emphasize this, a permutation of the integers 1, ..., n (denoted $k_1, ..., k_n$) will indicate the order of application.

Let $x^{(k_1)}$ be the vector among $x_1, ..., x_n$, whose image under the mapping $h_{k_1}(x)$ is the k_1th order statistic; i.e., $x^{(k_1)}$ is the observation x for which $k_1 - 1$ of $h_{k_1}(x)$ are less than $h_{k_1}(x^{(k_1)})$ and $n - k_1$ are larger. The *cutting function* h_{k_1} has an associated level set in $\mathbf{R}^p$ consisting of

$$\left\{ x \mid h_{k_1}(x) = h_{k_1}(x^{(k_1)}) \right\},$$

which defines a boundary between two *blocks*:

$$B_{1 \dots k_1} = \left\{ x \mid h_{k_1}(x) \le h_{k_1}(x^{(k_1)}) \right\},$$

$$B_{k_1+1 \dots n+1} = \left\{ x \mid h_{k_1}(x^{(k_1)}) < h_{k_1}(x) \right\}.$$

The union $B_{1 \dots k_1} \cup B_{k_1+1 \dots n+1} = \Omega$ (the sample space) and, in particular, $B_{1 \dots k_1}$ will contain exactly k_1 of the observations, and $B_{k_1+1 \dots n+1}$ will contain the remaining $n - k_1$ observations.

This process is continued, applying the functions $h_{k_1}(x), ..., h_{k_n}(x)$ in sequence to further subdivide $\mathbf{R}^p$ until, after n iterations, there remain $n + 1$ blocks $B_1, ..., B_{n+1}$ with $B_j \cap B_k = \phi$, $j \neq k$, and $\bigcup_i B_i = \Omega$. The function $h_{k_i}(x)$ that is applied at each stage, and the order of its application, is not chosen arbitrarily. It will be seen that the order of application is dictated by power considerations of an associated hypothesis test. To ensure that the ordering of $x_1, ..., x_n$ by $h_1(x), ..., h_n(x)$ is unique, excepting a set of measure zero, the requirement that $h_i(x)$ is continuous when x is distributed according to F(x) is imposed.

Before proceeding further, an illustrative example is appropriate (perhaps imperative).

Example 2.1. Consider the sample $\mathbf{X} = \{x_1, ..., x_n\}$, which is displayed as Table 1.

Table 1. Sample **X** (p = 2, n = 20)

i	$\mathbf{x}_i$	i	$\mathbf{x}_i$
1	(4.91, 2.16)	11	(4.21, 5.93)
2	(6.05, 5.54)	12	(0.15, 5.99)
3	(3.48, 1.35)	13	(9.31, 3.77)
4	(8.09, 0.18)	14	(4.10, 0.45)
5	(2.53, 3.49)	15	(5.83, 2.42)
6	(1.62, 2.46)	16	(6.00, 0.27)
7	(8.37, 2.29)	17	(3.30, 8.93)
8	(3.17, 6.27)	18	(4.38, 7.81)
9	(6.02, 4.51)	19	(4.93, 6.64)
10	(8.50, 4.65)	20	(1.22, 1.54)

A scatterplot of the data appears in Figure 1.

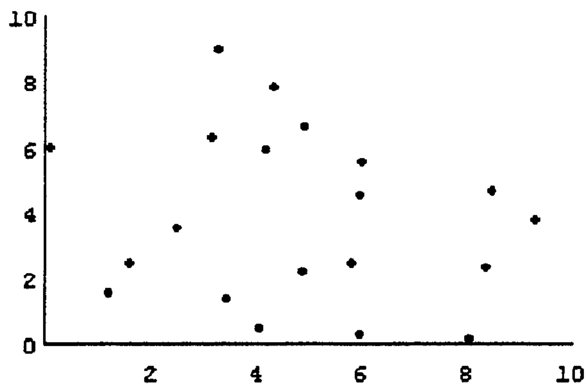

Figure 1. Scatterplot of sample **X**.

Arbitrary cutting functions $h_{k_1}(\mathbf{x}), \ldots, h_{k_{20}}(\mathbf{x})$, denoted

$$(c_2, c_2, c_1, c_2, c_2, c_2, c_1, c_2, c_2, c_1, c_1, c_1, c_2, c_1, c_2, c_1, c_2, c_1, c_1, c_2), \tag{2.1}$$

where

$$c_i(\mathbf{x}) = c_i((x_1, x_2)) = x_i, \quad i = 1, 2$$

and the corresponding permutation,

$$(10, 5, 15, 3, 7, 12, 18, 2, 4, 6, 8, 11, 13, 16, 19, 1, 9, 14, 17, 20), \tag{2.2}$$

will partition the sample space $\mathbf{R}^2$ into $n + 1 = 21$ blocks.

The first entry in the permutation (2.2) is $k_1 = 10$. The cutting function from (2.1), $h_{10} = c_1$, is applied to $\mathbf{x}_1, ..., \mathbf{x}_{20}$, and the pre-image of the 10th order statistic is determined to be $\mathbf{x}_{18} = (4.38, 7.81)$. This defines the first cut which divides the sample space into two parts (blocks):

$$B_{1 ... 10} = \{\mathbf{x} \mid c_1(\mathbf{x}) \leq 4.38\}$$

and

$$B_{11 ... 21} = \{\mathbf{x} \mid 4.38 < c_1(\mathbf{x})\} .$$

The second entry in the permutation is $k_2 = 5$. The cutting function $h_5 = c_2$ is applied next, but only to those sample points which are members of $B_{1...10}$ since the 5th order statistic will be bounded above by the 10th order statistic. The application of h_5 subdivides $B_{1...10}$ into $B_{1...5}$ and $B_{6...10}$. The next iteration, $k_3 = 15$ and $h_{15} = c_2$, paritions $B_{11...21}$ into $B_{11...15}$ and $B_{16...21}$ under the same argument; the 15th order statistic is bounded below by the 10th order statistic. These blocks are depicted in Figure 2.

This process is continued until each of the h_i has been applied. The sample space will be partitioned into 21 blocks, as depicted in Figure 3. In Figure 3, some representative blocks have been labelled. This example is referenced in following sections.

2.2 <u>Mathematical Foundation</u>. Thus far, discussion has been limited to the mechanics of block construction, without any motivation for engaging in such an exercise. Toward this end, consider

$$v_k = \int_{B_k} dF(\mathbf{x}), \quad k = 1, \ldots, n+1 .$$

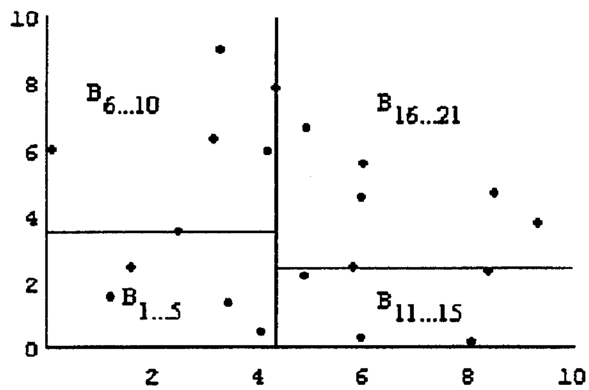

Figure 2. Non-negative $\mathbf{R}^2$ after cuts $h_{k_1}(x)$, $h_{k_2}(x)$, $h_{k_3}(x)$.

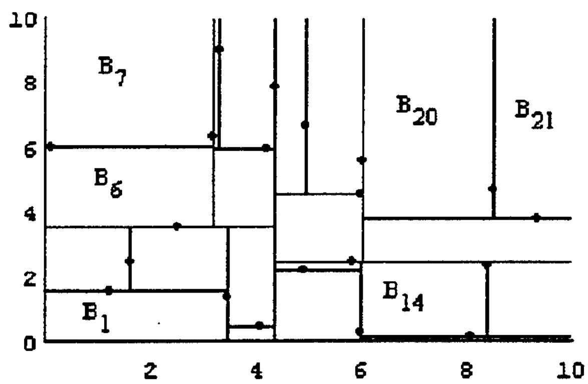

Figure 3. Non-negative $\mathbf{R}^2$ after cuts $h_{k_1}(x)$, ..., $h_{k_{20}}(x)$.

The *coverage* v_k is the probability assigned to block B_k under the distribution F(x). It can be shown (see Anderson [1966] and Wilks [1962], p. 238) that the coverages are distributed jointly as an n-variate Dirichlet distribution:

$$f(v_1, ..., v_n) = \frac{\Gamma(\eta_1 + ... + \eta_{n+1})}{\Gamma(\eta_1) \ ... \ \Gamma(\eta_{n+1})} v_1^{\eta_1 - 1} \ ... \ v_n^{\eta_n - 1} (1 - v_1 - ... - v_n)^{\eta_{n+1} - 1} . \tag{2.3}$$

The symmetry of the coverages v_1, ..., v_{n+1} in equation (2.3) leads to reference of the corresponding sample blocks B_1, ..., B_{n+1} as *statistically equivalent blocks.*

As a direct consequence of the manner in which the blocks are constructed, the expression

$$u^{(j)} = \sum_{k=1}^{j} v_k, \quad j = 1, ..., n$$

produces variates $u^{(1)}, ..., u^{(n)}$, which are distributed as the order statistics of a uniform random variable on [0, 1] (Anderson [1966]). The joint distribution of the variates is an ordered n-variate Dirichlet distribution (Wilks [1962], p. 236).

The importance of these results is twofold. First, the results do not depend upon the specific form of the distribution function F(x) and, as such, are distribution free. Second, testing for variates uniformly distributed may be accomplished through established procedures (see, for example, D'Agostino and Stephens [1986]).

3. MULTIVARIATE HYPOTHESIS TESTING

Statistically equivalent blocks find use in both one- and two-sample situations. In the one-sample case, a multivariate goodness-of-fit test may be accomplished. In the two-sample case, a test for identical distributions follows immediately from the procedure by which the blocks are formed.

3.1 One Sample. Section 2.2 provides the theoretical foundation for a multivariate goodness-of-fit procedure. Given a random sample $x_1, ..., x_n$ from an unknown distribution F(x), and a completely specified distribution G(x), the hypotheses

$$H_0: F(x) = G(x) \quad \forall x$$

and

$$H_1: F(x) \neq G(x)$$

may be established.

Example 3.1. Suppose that the data presented in Table 1 are to be tested for conformity to a bivariate uniform distribution on the square $[0, 10] \times [0, 10]$. The coverages of the blocks constructed in accordance with cutting functions (2.1) and permutation (2.2) under a bivariate uniform assumption are presented in Table 2. In the univariate case, the scalars $x_1, ..., x_n$ are naturally ordered $x^{(1)} < ... < x^{(n)}$, and a test of H_0:$F(x) = G(x)$ may be accomplished by determining whether $u^{(k)} = G(x^{(k)})$, $k = 1, \ldots, n$, are distributed uniformly on [0, 1].

Table 2. Block B_i Coverages

i	v_i	i	v_i
1	.0536	12	.0041
2	.0314	13	.0071
3	.0365	14	.0529
4	.0040	15	.0363
5	.0274	16	.0351
6	.0790	17	.0302
7	.1267	18	.0615
8	.0298	19	.0537
9	.0057	20	.1526
10	.0439	21	.0935
11	.0350		

The argument extends to the multivariate case, but the construction of an empirical cumulative distribution function for a random vector x does not hold as much intuitive appeal. The statistically equivalent blocks, once constructed from $x_1, \ldots, x_n$ as described in section 2.1, can be renumbered without loss of generality and accumulated to obtain alternative representations of a cumulative distribution function.

In consideration of this, a test based on probability assignment to intervals (blocks) without regard to location seems more appropriate. Fisher (1929) provides such a test in which the null hypothesis is rejected if

$$\Pr\{\max_j v_j > V\} = (N+1)(1-V)^N - {}_{N+1}C_2(1-2V)^N + \ldots + (-1)^{k-1}\,{}_{N+1}C_k(1-kV)^N$$

$$\frac{1}{k+1} \le V \le \frac{1}{k}, \; k=1, \ldots, N \tag{3.1}$$

exceeds a specified level of significance ε.

The test can be carried out by replacing V on the right side of (3.1) by $\max_j \nu_j$ and evaluating the expression. The computed value is the observed significance level, p. From Table 2, $\max_j \nu_j = \nu_{20} = 0.1526$; the observed significance level is $p = 0.63$, far too large to reject the null hypothesis of bivariate uniformity.

3.2 Two Sample. The decomposition of a p-dimensional sample space into statistically equivalent blocks allows for a ready extension to a two-sample test. Given independent random samples $\mathbf{X} = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ and $\mathbf{Y} = \{\mathbf{y}_1, ..., \mathbf{y}_m\}$ from unknown distributions F and G respectively, the hypotheses

$$H_0{:}F(\mathbf{x}) = G(\mathbf{x})\ \forall \mathbf{x}$$

and

$$H_1{:}F(\mathbf{x}) \neq G(\mathbf{x})$$

may be tested.

The mechanism for performing this test is perhaps more straightforward than for the one-sample test. The creation of the statistically equivalent blocks B_i, $i = 1, ..., n + 1$, imposes an ordering of the observations in $\mathbf{X}$ that was denoted by $\mathbf{x}^{(1)}, ..., \mathbf{x}^{(n)}$. Having created the blocks based on the sample $\mathbf{X}$, a relative ordering of the observations in $\mathbf{X}$ and $\mathbf{Y}$ denoted as "<<," according to the rule that $\mathbf{y}_i \in B_j$ iff $\mathbf{x}^{(j-1)} << \mathbf{y}_i << \mathbf{x}^{(j)}$, follows immediately. Under the null hypothesis, there should be no significant difference in the rank ordering of the observations from $\mathbf{X}$ (or $\mathbf{Y}$) in the combined sample. Therefore, any test based on relative ranking of the observations is appropriate for use in testing the hypothesis of identical distributions.

Example 3.2. Consider the sample $\mathbf{Y} = \{\mathbf{y}_1, ..., \mathbf{y}_m\}$ shown in Table 3.

Figure 4 displays the blocks which were created in section 2.1 based on the sample $\mathbf{X}$, with the points corresponding to sample $\mathbf{Y}$ overlaid. Based on the blocks into which the $\mathbf{Y}$ observations fall, the combined sample may be ordered as follows:

(x, x, x, x, x, x, x, x, x, x, x, y, x, y, x, y, y, y, x, y, y, y, y, y, x, x, x, x, y, y, y, y, y, x, y, y, x, y, y, y).

Table 3. Sample **Y** (p = 2, m = 20)

i	y_i	i	y_i
1	(13.90, 2.13)	11	(6.33, 4.44)
2	(7.71, 6.89)	12	(11.86, 0.83)
3	(9.67, 6.20)	13	(7.42, 2.31)
4	(7.56, 0.90)	14	(9.15, 3.94)
5	(10.39, 2.43)	15	(12.73, 6.12)
6	(13.47, 0.45)	16	(6.58, 3.04)
7	(14.55, 0.01)	17	(7.34, 3.69)
8	(7.46, 0.18)	18	(8.12, 2.59)
9	(11.25, 1.08)	19	(7.79, 3.68)
10	(13.0, 82.37)	20	(5.65, 2.40)

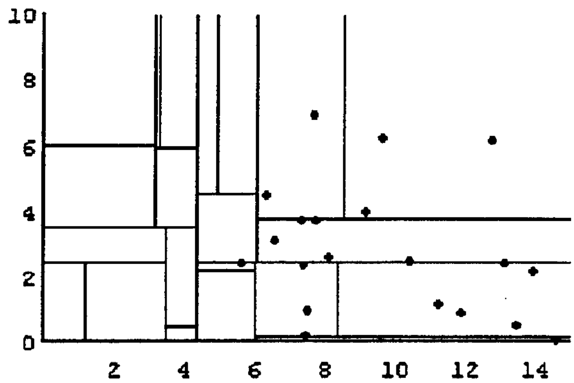

Figure 4. Blocks constructed from X with Y overlaid.

Any rank-based hypothesis test may be applied to this relative ordering. The Smirnov two-sample test provides a statistic of 0.55 and a corresponding p-value of 0.004. The Mann-Whitney test, which is primarily a test for difference in location, provides a p-value of less than 0.002.

4. PROXIMITY-BASED CUTTING FUNCTIONS

Cutting functions appearing in the literature are most often component-wise: $h(\mathbf{x}) = x_i$, a choice which facilitates presentation in two dimensions. Another class of cutting functions, based on proximity to the observations $\mathbf{X} = \{\mathbf{x}_i\}$ (where "proximity" is in an Euclidean metric sense), is now considered.

Let $\mathbf{z} \in \mathbf{R}^p$, the p-dimensional sample space containing the observations $\mathbf{X} = \{\mathbf{x}_i\}$ and $\mathbf{Y} = \{\mathbf{y}_j\}$. Consider a function $D:\mathbf{R}^p \to \mathbf{R}^n$ defined by $D_x(\mathbf{z}) = \{d_i\}$, where the d_i are the Euclidean distances from $\mathbf{z}$ to each of the observations $\mathbf{x}_i \in \mathbf{X}$. Without loss of generality, assume that $d_1 \leq d_2 \leq \ldots \leq d_n$. Now, for any real function $H:\mathbf{R}^n \to \mathbf{R}$ form the composite $h(\mathbf{z}) = H(D(\mathbf{z}))$. The function h will be called a *proximity-based cutting function* (PBCF) because the value taken on reflects the proximity of $\mathbf{z}$ to the members of $\mathbf{X}$.

Consider the expression

$$H(D(\mathbf{z})) = \sum_{i=1}^{n} \alpha_i d_i, \quad \alpha_i \geq 0. \tag{4.1}$$

It is clear that the order statistics $d^{(1)}, \ldots, d^{(m)}$, along with all linear combinations, are special cases of equation (4.1). In section 2.1, the requirement that $h_i(\mathbf{x})$ be continuous when $\mathbf{x}$ is distributed continuously was imposed. Since the distance functions d_i are clearly continuous, the expression (4.1) is also continuous, and the legitimacy of $h(\mathbf{z}) = H(D(\mathbf{z}))$ as a cutting function is established.

The motivation for examining this class of functions is as follows. In the two-sample case, the question "How closely does a random sample $\mathbf{X}$ resemble a random sample $\mathbf{Y}$?" is posed. Univariate rank tests address this problem following an argument that, under a null hypothesis of no difference, the sample $\mathbf{X}$ will be interspersed among the sample $\mathbf{Y}$. The choice of PBCF is an attempt to extend this argument to higher dimensions. Appropriately chosen PBCFs should partition the multidimensional space into statistically equivalent blocks that will distinguish when the observations under consideration are indeed in and among their counterpart.

Example 4.1. The choice of cutting functions in Example 2.1 gave level sets which were straight lines (or hyperplanes in higher dimension). The nature of the level sets and statistically equivalent blocks is not as intuitive for the PBCFs. Consider $H(D(z)) = d_2 + d_3$. This function maps a point $\mathbf{z} \in \mathbf{R}^p$ to the sum of the distances to the second and third closest $\mathbf{x} \in \mathbf{X}$. For the data from Example 2.1, this cutting function produces the level sets shown in Figure 5.

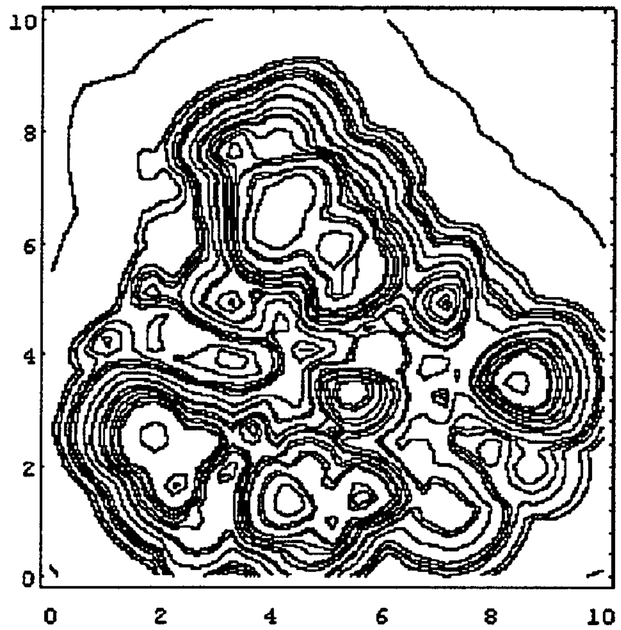

Figure 5. Blocks corresponding to $h(z) = d_2 + d_3$.

The statistically equivalent blocks do not resemble "blocks" at all for this choice of cutting function. Rather, the blocks are the areas bounded by level sets. This cutting function may be used to repeat the hypothesis test detailed in Example 3.2. In Figure 6, the sample **Y** has been overlaid on the blocks from Figure 5. Some of the level sets have been removed to allow the observations $\mathbf{y} \in \mathbf{Y}$ to be distinguished. A relative ordering of the two samples is again created. The Smirnov test yields a statistic of 0.45 for this ordering, which corresponds to a p-value of 0.034.

Since, in Example 3.2, the Smirnov test returned a p-value of 0.004, it would be unlikely to observe an even higher level of significance for these data and this hypothesis regardless of the choice of cutting function. In most practical situations, either value (0.004 or 0.034) is sufficient to abandon the null hypothesis. The intent is that PBCFs lead to a more powerful test of hypothesis, and while the notion that the sample **X** be interspersed among the sample **Y** under H_0 is not incorrect, it is incomplete. The requirement that **Y** be interspersed among the sample **X** is equally important.

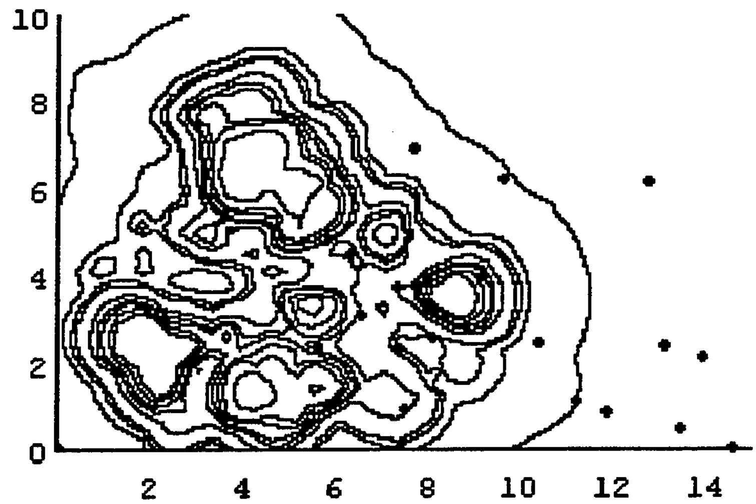

Figure 6. Sample X blocks with Y overlaid.

Consider the situation depicted in Figure 7. The level curves from Figure 5 have superimposed a *subset* of observations from Y that retain their integrity in the combined data set. Again, some level curves have been removed in order that the values from Y may be seen more clearly. The Smirnov statistic in this instance is 0.35, corresponding to a p-value of 0.264—the test has lost power against this type alternative.

The problem is that the mixture of x's and y's in the combined sample is not homogeneous. A direct approach to dealing with this situation is to reverse the roles of X and Y; i.e., construct blocks according to the sample Y and consider the dispersion of the sample X. The two tests of hypothesis can then be combined with a level of significance determined as follows. If the individual tests have significance levels α_1 and α_2, respectively, then the combined test has significance level $\alpha \leq \alpha_1 + \alpha_2$. To establish a level of significance α, it will suffice to set $\alpha_1 = \alpha_2 = \alpha/2$. If the individual tests have observed significance levels (a.k.a. critical levels) of p_1 and p_2, then the observed significance level for the combined test is $p = 2 \min (p_1, p_2)$.

Figure 8 illustrates blocks constructed from Y with X overlaid. The Smirnov statistic is 0.683, corresponding to a p-value of 0.0008. The critical level of the combined test procedure is then $p \leq 0.0016$.

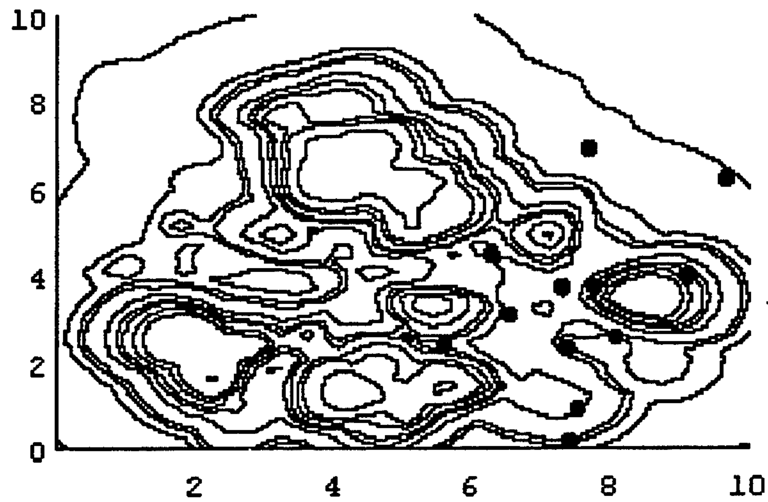

Figure 7. Sample X blocks with Y subset overlaid.

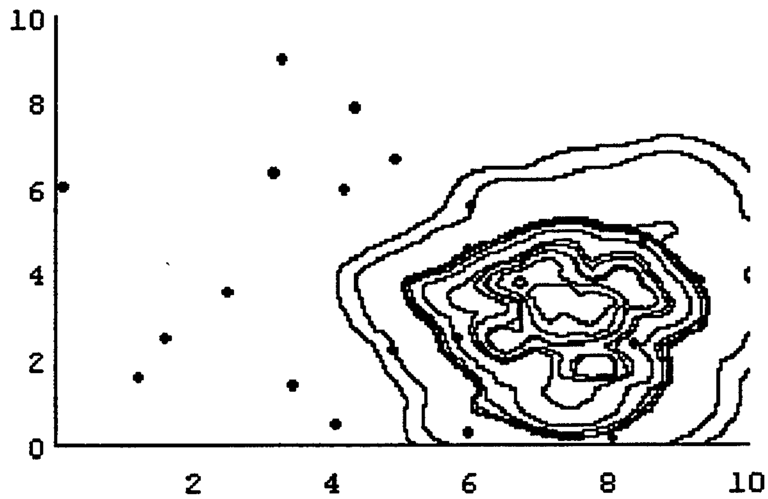

Figure 8. Blocks constructed from Y with X overlaid.

5. SUMMARY AND CONCLUSIONS

The sample **X**, introduced in Example 2.1, was taken from a uniform distribution on the square $[0,10] \times [0,10]$. Not surprisingly, the test of hypothesis of bivariate uniformity detailed in Example 3.1 produced an observed significance level $p = 0.63$, suggesting good agreement between data and hypothesis.

The sample **Y**, introduced in Example 3.2, was taken from a uniform distribution on the square $[5,15] \times [0,10]$. Again, both tests of hypotheses presented in Examples 3.2 and 4.1 detected the change in location even though the marginal distributions of the parent populations for **X** and **Y** coincide on the ordinate. The Mann-Whitney test appeared more sensitive to the shift in location.

The situation depicted in Figure 7 is one in which the level curves from Figure 5 have superimposed those observations from **Y** contained in $[5,10] \times [0,10]$. In an attempt to overcome an attendant loss of power, the roles of **X** and **Y** were interchanged and a combined test was performed. The combined test had an observed significance level of 0.0016.

The concept of PBCF holds promise for the analysis of multivariate data. Additional research is clearly in order. The power of the procedure has not been investigated; scaling of the variates in relation to the PBCF was not addressed; and only a single PBCF was illustrated. Computationally intensive methods for statistical data analysis is a natural extension of the powerful and economical computing resources that are readily available to the researcher and will continue to receive emphasis as a research area in mathematical statistics.

6. REFERENCES

Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: Wiley, 1958.

Anderson, T. W. "Some Nonparametric Multivariate Procedures Based on Statistically Equivalent Blocks." P. R. Krishnaiah (Editor), Multivariate Analysis, New York: Academic Press, pp. 5–27, 1966.

D'Agostino, R. B., and M. A. Stephens. Goodness-of-Fit Techniques. New York: Marcel Dekker, 1986.

Fisher, R. A. "Tests of Significance in Harmonic Analysis." Proceedings Royal Society London, Ser. A, vol. 125, pp. 54–59, 1929.

Fraser, D. A. S. Nonparametric Methods in Statistics. New York: Wiley, pp. 147–158, 1957.

Morrison, D. F. Multivariate Statistical Methods. New York: McGraw-Hill, 1976.

Puri, M. L., and P. K. Sen. Nonparametric Methods in Multivariate Analysis. New York: Wiley, 1971.

Wilks, S. S. Mathematical Statistics. New York: Wiley, 1962.

INTENTIONALLY LEFT BLANK.

NO. OF COPIES	ORGANIZATION
2	DEFENSE TECHNICAL INFO CTR ATTN DTIC DDA 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
1	DIRECTOR US ARMY RESEARCH LAB ATTN AMSRL OP SD TA 2800 POWDER MILL RD ADELPHI MD 20783-1145
3	DIRECTOR US ARMY RESEARCH LAB ATTN AMSRL OP SD TL 2800 POWDER MILL RD ADELPHI MD 20783-1145
1	DIRECTOR US ARMY RESEARCH LAB ATTN AMSRL OP SD TP 2800 POWDER MILL RD ADELPHI MD 20783-1145
	ABERDEEN PROVING GROUND
5	DIR USARL ATTN AMSRL OP AP L (305)

NO. OF COPIES	ORGANIZATION
1	SUPERINTENDANT US MILITARY ACADEMY ATTN MADN A WEST POINT NY 10996-1786
5	SAA ATTN MAJ KEITH E MATTHEWS SUITE 404 5111 LEESBURG PIKE FALLS CHURCH VA 22041-3210
1	US MILITARY ACADEMY DEPT OF MATH SCI ATTN MAJ DON ENGEN WEST POINT NY 10996-1786
1	US MILITARY ACADEMY DEPT OF MATH SCI ATTN CPT ROD STURDIVANT WEST POINT NY 10996-1786
1	US ARMY RESEARCH OFFICE MATH & COMP SCI DIV ATTN DR ROBERT R LAUNER PO BOX 12211 RESEARCH TRIANGLE PARK NC 27709-2211
1	RICE UNIVERSITY ATTN PROF JAMES R THOMPSON CHAIRMAN DEPT OF STATISTICS HOUSTON TX 77251-1892
1	WALTER REED ARMY INST OF RSRCH WALTER REED ARMY MED CTR ATTN DOUGLAS B TANG PH D CHIEF DEPT OF BIOSTATISTICS WASHINGTON DC 20307-5100
1	DIRECTOR US ARMY RESEARCH LABORATORY ATTN AMSRL SC I J GANTT 115 OKEEFE BLDG GA INST OF TECH ATLANTA GA 30332-0800

NO. OF COPIES	ORGANIZATION
	ABERDEEN PROVING GROUND
17	DIR, USARL ATTN: AMSRL-SC, W. H. MERMAGEN AMSRL-SC-S, A. MARK M.S. TAYLOR (7 CP) B. BODT AMSRL-SC-CC, A. CELMINS C. ZOLTANI AMSRL-SC-II, J. DUMER T. HANRATTY AMSRL-SL-BV, W. BAKER AMSRL-WT-PB, D. WEBB AMSRL-IS-TP, A. BRODEEN

USER EVALUATION SHEET/CHANGE OF ADDRESS

This Laboratory undertakes a continuing effort to improve the quality of the reports it publishes. Your comments/answers to the items/questions below will aid us in our efforts.

1. ARL Report Number ARL-TR-992 Date of Report April 1996

2. Date Report Received ________________

3. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which the report will be used.) ________________

4. Specifically, how is the report being used? (Information source, design data, procedure, source of ideas, etc.) ________________

5. Has the information in this report led to any quantitative savings as far as man-hours or dollars saved, operating costs avoided, or efficiencies achieved, etc? If so, please elaborate. ________________

6. General Comments. What do you think should be changed to improve future reports? (Indicate changes to organization, technical content, format, etc.) ________________

CURRENT ADDRESS

Organization

Name

Street or P.O. Box No.

City, State, Zip Code

7. If indicating a Change of Address or Address Correction, please provide the Current or Correct address above and the Old or Incorrect address below.

OLD ADDRESS

Organization

Name

Street or P.O. Box No.

City, State, Zip Code

(Remove this sheet, fold as indicated, tape closed, and mail.)

(DO NOT STAPLE)